

# A Note on Limit Theorems for Multivariate Martingales\*

Uwe Küchler  
Humboldt Universität zu Berlin

Michael Sørensen  
University of Copenhagen

## Abstract

Multivariate versions of the law of large numbers and the central limit theorem for martingales are given in a generality that is often necessary when studying statistical inference for stochastic process models. To illustrate the usefulness of the results, we consider estimation for a multi-dimensional Gaussian diffusion, where results on consistency and asymptotic normality of the maximum likelihood estimator are obtained in cases that were not covered by previously published limit theorems. The results are also applied to martingales of a different nature, which are typical of the problems occurring in connection with statistical inference for stochastic delay equations.

**Key words:** Central limit theorem; multivariate Gaussian diffusions; likelihood inference; stochastic delay equations; weak law of large numbers.

---

\* Research for this paper was partially carried out within Sonderforschungsbereich 373. This paper was printed using funds made available by the Deutsche Forschungsgemeinschaft.

# 1 Introduction

The law of large numbers and the central limit theorem for martingales have proved very useful tools for obtaining asymptotic results about estimators of parameters in stochastic process models. However, the multivariate versions of these results published so far have been stated either with conditions that are too strict for many statistical applications, or in a generality that makes application to particular statistical models very difficult. In this paper we give multivariate versions of the law of large numbers and the central limit theorem for martingales in a generality that seems to us useful in many statistical applications.

When studying inference for stochastic process models where the data is a single trajectory, a central limit theorem for martingales is often a useful tool for obtaining asymptotic distributional results as the length of the observation period goes to infinity, see e.g. Barndorff-Nielsen and Sørensen (1991, 1994). What is needed is a central limit theorem for a properly normalized martingale as the time goes to infinity. When the statistical parameter is multidimensional or of infinite dimension, a multivariate central limit theorem is needed. We briefly review some previously published multivariate central limit theorems, which have proved useful in the type of statistical application indicated above when the models may be non-ergodic. For a review of older work on mainly one-dimensional martingales, see Helland (1982). A multidimensional martingale central limit theorem applicable to some non-ergodic models was given by Hutton and Nelson (1984). However, their assumption that the quadratic variation matrix converges when normalized by a scalar is rather restrictive. In order to be able to obtain results for multivariate martingales, the coordinates of which increase at different rates, a type of martingales which occurs quite naturally when analysing many statistical stochastic process models, Sørensen (1991) published a result, where the quadratic variation matrix converges when normalized by a diagonal matrix. This has also turned out to be too restrictive because, for instance, it does not cover the case where all components of the martingale increase at the same rate, but where there are directions different from the main axes in which the rate of increase is smaller. This happens in some stochastic process models, see Dietz (1992), Stockmarr (1996) and Gushchin and Küchler (1997). In Section 2 we give a central limit theorem for multivariate mar-

tingales, where the quadratic variation matrix is assumed to converge when normalized by a suitable full matrix. This more general result covers also martingales of the type just described. In Section 3 it is demonstrated how it can be used to find the asymptotic distribution of the maximum likelihood estimator in multivariate Gaussian diffusion models in cases that could not be treated by the previous, less general, central limit theorems. In Section 4 the central limit theorem is applied to a type of martingales that occur in connection with investigations of statistical inference for solutions of stochastic delay equation.

Laws of large numbers for martingales are useful tools for proving consistency of estimators in the type of statistical models described in the previous paragraph. These results state that, as time tends to infinity, a martingale normalized by its quadratic variation converges to zero almost surely or in probability on the set where the quadratic variation tends to infinity. A first result was published by Lépingle (1978), and the paper by Liptser (1980) clarified the situation for one-dimensional martingales. The strong law of large numbers for martingales is proved by means of a generalization of the Kronecker lemma. Melnikov (1986) was able to prove a multivariate version of the strong law of large numbers for martingales under the assumption that the limsup of the ratio of the largest to the smallest eigenvalue of the quadratic variation matrix is finite. This assumption is rather strong, and in several statistical applications Melnikov's result can not be applied. Melnikov's condition was considerably weakened in the laws of large number for martingales by Le Breton and Musiela (1986, 1989) and Kaufmann (1987), but their conditions are still too strong to cover cases where some eigenvalues of the quadratic variation increase linearly while others increase exponentially, which can happen even in relatively simple statistical models. Le Breton and Musiela (1987) and Dzhaparidze and Spreij (1993) proved a strong law of large numbers for multivariate martingales which are Gaussian or which have deterministic quadratic variation, respectively. In these two papers the only condition is, essentially, that the inverse of the quadratic variation matrix tends to zero. The assumptions that the martingale is Gaussian or that the quadratic variation matrix is deterministic are also too restrictive in many applications. In Section 2 a *weak* law of large numbers for multivariate martingales is derived as a corollary to the general central limit theorem. The result holds without any further conditions, so by proving only conver-

gence in probability, we avoid the restrictive conditions in the strong laws of large numbers for multivariate martingales published so far and obtain a more broadly applicable result. Apart from the applications mentioned above, the weak law of large numbers for multivariate martingales can be used to prove consistency of least squares estimators in a very general class of semimartingale models; see Melnikov and Novikov (1990).

## 2 Limit Theorems

Let  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$  be a complete filtered probability space. Let  $M = (M_1, \dots, M_k)^T$  be a  $k$ -dimensional square integrable martingale with respect to  $\{\mathcal{F}_t\}$ , i.e. the coordinates of  $M_t$  are square integrable for all  $t \geq 0$ . We can therefore for all  $t \geq 0$  define  $H_t = E(M_t M_t^T)$ , where  $T$  denotes transposition. We assume that the sample paths of  $M$  are right continuous with limits from the left. Let  $[M]_t$  be the quadratic variation matrix of  $M$ , i.e. the  $(i, j)$ 'th entry of  $[M]_t$  is  $[M_i, M_j]_t$ . By  $I_k$  we denote the  $k \times k$  identity matrix, and by  $\det(A)$  the determinant of a square matrix  $A$ , while  $A^{\frac{1}{2}}$  denotes the unique positive semi-definite square root of a positive semi-definite matrix  $A$ . The concepts of stable and mixing convergence used in the following theorem were introduced by Rényi (1958, 1963) and further developed by Aldous and Eagleson (1978), see also the discussion in Hall and Heyde (1980).

**Theorem 2.1** *Suppose there exists a family of invertible non-random  $k \times k$ -matrices  $\{K_t : t > 0\}$ , with  $t \mapsto K_t$  continuous, such that as  $t \rightarrow \infty$*

- a)  $K_t \rightarrow 0$ ,
- b)  $\bar{K}_{it} E(\sup_{s \leq t} |\Delta M_{is}|) \rightarrow 0, i = 1, \dots, k$ ,  
where  $\bar{K}_{it} = \sum_{j=1}^k |K_{jit}|$  and  $\Delta M_{is} = M_{is} - M_{is-}$ ,
- c)  $K_t [M]_t K_t^T \rightarrow \eta^2$  in probability,  
where  $\eta^2$  is a random positive semi-definite matrix, and
- d)  $K_t H_t K_t^T \rightarrow \Sigma$ ,  
where  $\Sigma$  is a positive definite matrix.

*Then we have the following results on convergence in distribution as  $t \rightarrow \infty$ :*

$$K_t M_t \rightarrow Z, \quad (\text{stably}) \quad (2.1)$$

where the distribution of  $Z$  equals that of  $\eta^2 U$ , where  $U$  is a standard normal distributed  $k$ -dimensional random vector independent of  $\eta^2$ . This distribution is the normal variance-mixture with characteristic function  $\varphi(u) = E(\exp[-\frac{1}{2}u^T \eta^2 u])$ ,  $u = (u_1, \dots, u_k)^T$ . Moreover,

$$(K_t M_t, K_t [M]_t K_t^T) \rightarrow (\eta^2 U, \eta^2). \quad (2.2)$$

Provided that  $P(\det(\eta^2) > 0) > 0$ , we have the following results on convergence in distribution conditional on  $\{\det(\eta^2) > 0\}$ :

$$(K_t [M]_t K_t^T)^{-\frac{1}{2}} K_t M_t | \{\det(\eta^2) > 0\} \rightarrow N(0, I_k) \quad (\text{mixing}) \quad (2.3)$$

and

$$M_t^T [M]_t^{-1} M_t | \{\det(\eta^2) > 0\} \rightarrow \chi^2(k) \quad (\text{mixing}). \quad (2.4)$$

**Proof:** To prove (2.1) it suffices to show that  $x^T K_t M_t \rightarrow x^T Z$  (stably) for all  $x \in \mathbb{R}^k \setminus \{0\}$  (the Cramér-Wold device). For every  $t > 0$  the process

$$X_s^t = (x^T K_t H_t K_t^T x)^{-\frac{1}{2}} x^T K_t M_{st}, \quad 0 \leq s \leq 1,$$

is a one-dimensional square integrable martingale with respect to the filtration  $\{\mathcal{F}_{st}\}$ . For  $t$  large enough  $x^T K_t H_t K_t^T x > 0$  because  $\Sigma$  is positive definite. Using b) and d), we see that

$$\begin{aligned} E(\sup_{0 \leq s \leq 1} |\Delta X_s^t|) &\leq \\ &(x^T K_t H_t K_t^T x)^{-\frac{1}{2}} \max_i |x_i| \sum_j \bar{K}_{jt} E(\sup_{s \leq t} |\Delta M_{js}|) \rightarrow 0 \end{aligned}$$

as  $t \rightarrow \infty$ . From c) and d) it follows that

$$[X^t]_1 = (x^T K_t H_t K_t^T x)^{-1} x^T K_t [M]_t K_t^T x \rightarrow (x^T \Sigma x)^{-1} x^T \eta^2 x$$

in probability as  $t \rightarrow \infty$ . Thus the class of processes  $X^t, t > 0$ , satisfies the same conditions as the martingales (3.6) in the proof of Theorem 2 in Feigin

(1985). Therefore, it follows from Feigin's proof that  $X_1^t$  converges stably in distribution to the zero-mean normal variance-mixture with the distribution of  $x^T \eta^2 x$  as mixing distribution, i.e. to the distribution of  $x^T Z$ .

The stability of (2.1) implies (2.2) and (2.3), see Aldous and Eagleson (1978). Remember that on  $\{\det(\eta^2) > 0\}$  the matrix  $K_t[M]_t K_t^T$  is positive definite for  $t$  large enough. The result (2.4) follows immediately from (2.3).  $\square$

In applications a main problem is to find the family of matrices  $\{K_t : t > 0\}$ . Based on c) and d) in Theorem 2.1 one might try to get an idea of how to choose  $K_t$  by studying the rate of increase of the entries of  $H_t$ , but this is not always enough, because these entries may all increase at the same rate even when  $M$  grows at different rates in directions that are not parallel to the coordinate axes. Often one must search for a family  $\{K_t : t > 0\}$  of the form  $D_t C$  where  $D_t$  is a diagonal matrix, while  $C$  changes the coordinate axes appropriately. We shall give examples in the Sections 3 and 4.

Since  $H_t$  is positive semi-definite, there exists an orthogonal matrix  $O_t$  and a diagonal matrix  $D_t$  with non-negative diagonal elements, such that  $H_t = O_t^T D_t O_t$ . If we can use  $K_t = D_t^{-\frac{1}{2}} O_t$  or  $K_t = O_t^T D_t^{-\frac{1}{2}} O_t$ , the condition d) in Theorem 2.1 is automatically satisfied.

The following weak law of large numbers follows immediately from Theorem 2.1.

**Corollary 2.2** *Assume the conditions of Theorem 2.1. Then  $[M]_t$  is invertible on  $\{\det(\eta^2) > 0\}$  for  $t$  sufficiently large, and*

$$[M]_t^{-1} M_t \rightarrow 0 \tag{2.5}$$

*in probability on  $\{\det(\eta^2) > 0\}$  as  $t \rightarrow \infty$ .*

**Proof:** The result follows from Theorem 2.1 because

$$[M]_t^{-1} M_t = K_t^T (K_t [M]_t K_t^T)^{-1} K_t M_t,$$

where  $K_t M_t$  converges in distribution and hence is stochastically bounded,  $(K_t [M]_t K_t^T)^{-1} \rightarrow (\eta^2)^{-1}$  in probability on  $\{\det(\eta^2) > 0\}$ , and  $K_t \rightarrow 0$ .  $\square$

### 3 Multi-dimensional Gaussian Diffusions

In this section we consider the class of  $k$ -dimensional Gaussian diffusions given as solutions to the stochastic differential equation

$$dX_t = BX_t dt + dW_t, \quad X_0 = 0, \quad (3.1)$$

where  $X_t = (X_{1t}, \dots, X_{kt})^T$ ,  $B$  is a  $k \times k$ -matrix and  $W$  is a  $k$ -dimensional standard Wiener process.

If  $\theta = (b_{11}, \dots, b_{1k}, \dots, b_{k1}, \dots, b_{kk})^T$ , then the likelihood function based on observation of  $X_s$  for  $s \in [0, t]$  is

$$L_t(\theta) = \exp(N_t \theta - \frac{1}{2} \theta^T I_t \theta), \quad (3.2)$$

where  $N_t$  is the  $k^2$ -dimensional vector

$$\left( \int_0^t X_{1s} dX_{1s}, \dots, \int_0^t X_{ks} dX_{1s}, \dots, \int_0^t X_{1s} dX_{ks}, \dots, \int_0^t X_{ks} dX_{ks} \right)^T$$

and  $I_t$  is the  $k^2 \times k^2$ -matrix

$$I_t = \begin{pmatrix} J_t & 0 & \cdots & 0 \\ 0 & J_t & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & J_t \end{pmatrix},$$

where  $J_t$  is the  $k \times k$ -matrix  $\int_0^t X_s X_s^T ds$ , which is almost surely invertible. The maximum likelihood estimator is  $\hat{\theta}_t = I_t^{-1} N_t$ . Note that  $\hat{\theta}_t = I_t^{-1} (I_t \theta + M_t) = \theta + I_t^{-1} M_t$ , where  $M_t$  is the  $k^2$ -dimensional square integrable martingale  $M_t = (\int_0^t X_{1s} dW_{1s}, \dots, \int_0^t X_{ks} dW_{1s}, \dots, \int_0^t X_{1s} dW_{ks}, \dots, \int_0^t X_{ks} dW_{ks})$ . Since  $[M]_t = I_t$ , the consistency of  $\hat{\theta}$  follows if we can apply Corollary 2.2 to  $M$ . Moreover, by Theorem 2.1, also the asymptotic normality of  $(K_t I_t K_t^T)^{\frac{1}{2}} (K_t^{-1})^T (\hat{\theta}_t - \theta) = (K_t I_t K_t^T)^{-\frac{1}{2}} K_t M_t$  (or more simply that  $(\hat{\theta}_t - \theta)^T I_t (\hat{\theta}_t - \theta) = M_t^T I_t^{-1} M_t \rightarrow \chi^2(k^2)$ ) follows too. We need to find  $K_t$  such that c) and d) of Theorem 2.1 are satisfied.

The solution to (3.1) is

$$X_t = \int_0^t e^{B(t-s)} dW_s. \quad (3.3)$$

Suppose  $B$  can be diagonalized, i.e. that there exists an invertible matrix  $C$  and a diagonal matrix  $D$  such that  $B = C^{-1}DC$ . Since  $B^n = C^{-1}D^nC$ , we find that

$$CX_t = \int_0^t e^{D(t-s)} d\tilde{W}_s, \quad (3.4)$$

where  $\tilde{W}_s = CW_s$  is a  $k$ -dimensional Wiener process with coordinate processes that are typically not independent. In fact,  $\langle \tilde{W} \rangle_t = tCC^T$ . Hence  $(i, j)$ 'th entry of

$$CJ_tC^T = \int_0^t CX_sX_s^TC^Tds. \quad (3.5)$$

is

$$a_{ij}(t) = \int_0^t \int_0^s e^{d_i(s-u)} d\tilde{W}_{iu} \int_0^s e^{d_j(s-u)} d\tilde{W}_{ju} ds, \quad (3.6)$$

where  $d_i, i = 1, \dots, k$  are the diagonal elements of  $D$ , i.e. the eigenvalues of  $B$ . If  $d_i > 0$  and  $d_j > 0$ , the Toeplitz lemma implies that

$$e^{-t(d_i+d_j)} a_{ij}(t) \rightarrow (d_i + d_j)^{-1} \int_0^\infty e^{-d_i u} d\tilde{W}_{iu} \int_0^\infty e^{-d_j u} d\tilde{W}_{ju}$$

almost surely as  $t \rightarrow \infty$ . If  $d_i < 0$  and  $d_j < 0$ , then  $t^{-1}a_{ij}(t) \rightarrow -\zeta_{ij}/(d_i + d_j)$  almost surely as  $t \rightarrow \infty$ , where  $\zeta_{ij}$  is the  $(i, j)$ 'th entry of  $CC^T$ . This is because the two processes under the Lebesgue integral in (3.6) are in this case ergodic. Finally, suppose  $d_i > 0$  and  $d_j < 0$ . Then  $t^{-\frac{1}{2}}e^{-d_i t}a_{ij}(t) \rightarrow 0$  in probability as  $t \rightarrow \infty$  by a generalization of the Toeplitz lemma because  $\int_0^t e^{d_j(t-s)} d\tilde{W}_{js}$  converges in distribution and hence is stochastically bounded.

Thus if all eigenvalues of  $B$  are real and different from zero, we can define  $K_t$  by

$$K_t = \begin{pmatrix} A_t C & 0 & \cdots & 0 \\ 0 & A_t C & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & A_t C \end{pmatrix},$$



where  $A_t$  is the diagonal matrix  $A_t = \text{diag}(\varphi_1(t), \dots, \varphi_k(t))$ , with

$$\varphi_i(t) = \begin{cases} e^{-d_i t} & \text{if } d_i > 0 \\ t^{-\frac{1}{2}} & \text{if } d_i < 0. \end{cases}$$

We have shown that  $K_t I_t K_t^T \rightarrow \eta^2(\theta)$  in probability as  $t \rightarrow \infty$ , where  $\eta^2(\theta)$  is random if one or more of the eigenvalues of  $B$  are positive. It is not a restriction of the generality to assume that  $d_i < 0$  for  $i = 1, \dots, m$  and  $d_i > 0$  for  $i = m + 1, \dots, k$ , where  $0 \leq m \leq k$ . Then  $\eta^2$  has the form

$$\eta^2 = \begin{pmatrix} \eta_1^2 & O \\ O^T & \eta_2^2 \end{pmatrix},$$

where  $O$  is an  $m \times (k - m)$ -matrix with all entries equal to zero. To prove that  $\eta^2$  is positive definite it is enough to prove that each of the two matrices  $\eta_1^2$  and  $\eta_2^2$  is positive definite. Obviously, they are positive semi-definite, so it is enough to prove that they are invertible. To impose necessary conditions for this to hold we need the concept of controllability. Let  $R$  and  $V$  be  $d \times d$ -matrices, where  $V$  is positive semi-definite, then  $(R, V)$  are called controllable if the rank of the  $d \times d^2$ -matrix  $[V, RV, \dots, R^{d-1}V]$  is  $d$ . We need the following result, which is well-known in control theory. For a proof see Davis (1977, Chapter 4) or Chaleyat-Maurel and Elie (1981, Lemma 1.10).

**Lemma 3.1** *The matrix  $\int_0^t \exp(sR)V \exp(sR^T)ds$  is invertible (i.e. positive definite) for  $t > 0$  if and only if  $(R, V)$  are controllable.*

Now,  $\eta_1^2 = \int_0^\infty \exp(sR_1)V_1 \exp(sR_1)ds$ , where  $R = \text{diag}(d_1, \dots, d_m)$  and  $V_1$  is the upper left  $m \times m$  submatrix of  $CC^T$ , so  $\eta_1^2$  is positive definite if  $(R_1, V_1)$  are controllable. This is, for instance, the case if  $V_1$  is invertible. The random matrix  $\eta_2^2$  is positive definite if the matrix  $\delta$  with entries  $\delta_{ij} = (d_i + d_j)^{-1}$  is positive definite. Since  $\delta = \int_0^\infty \exp(sR_2)V_2 \exp(sR_2)ds$ , where  $R_2 = \text{diag}(-d_{m+1}, \dots, -d_k)$  and  $V_2$  is the  $(k - m) \times (k - m)$ -matrix with all entries equal to one, we see that  $\eta_2^2$  is positive definite if  $(R_2, V_2)$  are controllable. This is the case if and only if the eigenvalues  $d_{m+1}, \dots, d_k$  are all different.

We have now given conditions ensuring that c) of Theorem 2.1 is satisfied. Let us turn to condition d) in the same theorem. Since

$$E(a_{ij}(t)) = \zeta_{ij}(d_i + d_j)^{-2} \left\{ e^{(d_i + d_j)t} - 1 - t(d_i + d_j) \right\},$$

we see that  $A_t C E_\theta(J_t) C^T A_t \rightarrow \Sigma(\theta)$ , where

$$\Sigma_{ij}(\theta) = \begin{cases} \zeta_{ij}(d_i + d_j)^{-2} & \text{if } d_i > 0 \text{ and } d_j > 0 \\ -\zeta_{ij}(d_i + d_j)^{-1} & \text{if } d_i < 0 \text{ and } d_j < 0 \\ 0 & \text{if } d_i > 0 \text{ and } d_j < 0. \end{cases}$$

As  $H_t = E_\theta(I_t)$ , we see that we just have to check that  $\Sigma$  is positive definite. Under the assumption made earlier, that  $d_i < 0$  for  $i = 1, \dots, m$  and  $d_i > 0$  for  $i = m + 1, \dots, k$ , where  $0 \leq m \leq k$ , the matrix  $\Sigma$  is of the form

$$\Sigma = \begin{pmatrix} \Sigma_{11} & O \\ O^T & \Sigma_{22} \end{pmatrix}.$$

We have already seen that  $\Sigma_{11} = \eta_1^2$  is positive definite if  $(R_1, V_1)$  is controllable. Note that  $\Sigma_{22} = \int_0^\infty \exp(sR_2)V_3 \exp(sR_2)ds$ , where the  $(i, j)$ 'th entry of the  $(k - m) \times (k - m)$  matrix  $V_3$  is  $\zeta_{ij}/(d_i + d_j)$ ,  $m + 1 \leq i, j \leq k$ . If  $V_3$  is invertible,  $(R_2, V_3)$  are controllable, so that by Lemma 3.1 the matrix  $\Sigma_{22}$  is positive definite. That  $V_3$  is invertible follows in the same way as the invertibility of  $\eta_1^2$ , if we assume that  $(R_2, V_4)$  are controllable, where  $V_4$  is the lower right  $(k - m) \times (k - m)$  submatrix of  $CC^T$ . In particular, we can assume that  $V_4$  is invertible.

We have thus proved consistency and, after normalization as described above, asymptotic normality of the maximum likelihood estimator provided that all eigenvalues of  $B$  are real and different from zero, that all positive eigenvalues are different, and that  $CC^T$  is positive definite.

This could not have been done for all these values of  $B$  by means of e.g. the martingale limit theorem in Sørensen (1991). Consider for instance the 2-dimensional diffusion with

$$B = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}.$$

Then  $B^2$  equals the identity matrix, so  $e^{Bu} = I \cosh u + B \sinh u$ . Hence

$$X_t = \begin{pmatrix} \int_0^t e^{(t-s)} dW_{1s} \\ \int_0^t \sinh(t-s) dW_{1s} + \int_0^t e^{-(t-s)} dW_{2s} \end{pmatrix}$$

It is not difficult to see that all entries of  $E(J_t)$  are of order  $e^{2t}$ , and that

$$e^{-2t}E(J_t) \rightarrow \begin{pmatrix} \frac{1}{4} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{16} \end{pmatrix},$$

which is singular, so there is no way of normalizing  $J_t$  and  $E(J_t)$  with a diagonal matrix to obtain a non-singular limit. However,  $B$  can be diagonalized:

$$\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix},$$

so the result proved above holds.

Maximum likelihood estimation for multivariate Gaussian diffusions has been studied by several authors, see Le Breton (1977, 1984), Le Breton and Musiela (1982, 1985) and Stockmarr (1996). None of these authors proved consistency or asymptotic normality in the case where some eigenvalues of  $B$  are positive while others are negative.

## 4 Martingales with time-delay.

In this section, we consider limit results for martingales of type that is typical of problems occurring in the study of statistical inference for stochastic delay equations.

Assume that  $M^{(1)}$  and  $M^{(2)}$  are continuous martingales with  $\langle M^{(i)}, M^{(j)} \rangle_t = \alpha(t)$ , where  $\alpha(t)$  is independent of  $i, j \in \{1, 2\}$  and differentiable with respect to  $t$ . We suppose further that  $\alpha'(t)$  has an almost surely finite limit, which we denote by  $\alpha'(\infty)$ . Define

$$Y_i(t) = \int_0^t x_0(t-s) dM_s^{(i)}, \quad t \geq 0, \quad i = 1, 2,$$

where

$$x_0(u) = e^{\lambda_1 u} + e^{\lambda_2 u}, \quad u \in \mathbb{R},$$

with  $0 < \lambda_2 < \lambda_1$  fixed. Then it holds  $P$ -almost surely that

$$e^{-\lambda_1 t} Y_i(t) \rightarrow U_i = \int_0^\infty e^{-\lambda_1 s} dM^{(i)}(s), \quad i = 1, 2,$$

as  $t \rightarrow \infty$ .

For  $t \geq 0$ , define the two-dimensional martingale

$$M_t = \begin{pmatrix} \int_0^t Y_1(s) dM^{(1)}(s) \\ \int_0^t Y_2(s-1) dM^{(2)}(s) \end{pmatrix},$$

which has the quadratic variation matrix

$$[M]_t = \begin{pmatrix} \int_0^t Y_1^2(s) d\alpha(s) & \int_0^t Y_1(s) Y_2(s-1) d\alpha(s) \\ \int_0^t Y_1(s) Y_2(s-1) d\alpha(s) & \int_0^t Y_2^2(s-1) d\alpha(s) \end{pmatrix}.$$

Elementary calculations show that, after normalization by  $\exp(-2\lambda_1 t)$ , all entries of  $[M]_t$  tend to (random) finite limits as  $t \rightarrow \infty$ ,

$$\exp(-2\lambda_1 t)[M]_t \rightarrow \frac{(\alpha'(\infty))^2}{2\lambda_1} \begin{pmatrix} 1 & e^{-\lambda_1} \\ e^{-\lambda_1} & e^{-2\lambda_1} \end{pmatrix},$$

and that the limit matrix is singular. This can be seen as the normalization  $K_t[M]_t K_t$  with  $K_t = e^{-\lambda_1 t} I_2$ . No other normalization of  $[M]_t$  by deterministic diagonal matrices, possibly with different rates of increase for the diagonal elements, lead to a regular limit matrix either.

It is, however, possible to choose a non-diagonal matrix  $K_t$  for which a non-singular limit matrix is obtained. In fact, for  $K_t = \psi(t)\Phi$  with

$$\Phi = \begin{pmatrix} 1 & 0 \\ 1 & -e^{\lambda_1} \end{pmatrix} \quad \text{and} \quad \psi(t) = \begin{pmatrix} e^{-\lambda_1 t} & 0 \\ 0 & e^{-\lambda_2 t} \end{pmatrix},$$

the expression  $K_t[M]_t K_t^T$  tends to

$$\eta^2 = (\alpha'(\infty))^2 \begin{pmatrix} \frac{U_1^2}{2\lambda_1} & \frac{U_1 U_2}{\lambda_1 + \lambda_2} (1 - e^{(\lambda_1 - \lambda_2)}) \\ \frac{U_1 U_2}{\lambda_1 + \lambda_2} (1 - e^{(\lambda_1 - \lambda_2)}) & \frac{U_2^2}{2\lambda_2} (1 - e^{(\lambda_1 - \lambda_2)})^2 \end{pmatrix},$$

which is non-singular with probability one. To see this, note that  $\eta^2$  has the structure  $U\Lambda U$ , where  $U = \text{diag}(U_1, U_2)$  and where  $\Lambda$  is seen to be positive definite by arguments similar to those given in the previous section.

## References

- Aldous, D.J. and Eagleson, G.K. (1978): On mixing and stability of limit theorems. *Ann. Probability* **6**, 325–331.
- Barndorff-Nielsen, O.E. and Sørensen, M. (1991): Information quantities in non-classical settings. *Computational Statistics and Data Analysis* **12**, 143–158.
- Barndorff-Nielsen, O.E. and Sørensen, M. (1994): A review of some aspects of asymptotic likelihood theory for stochastic processes. *Internat. Statist. Rev.* **62**, 133–165.
- Chaleyat-Maurel, M. and Elie L. (1981): Diffusions Gaussiennes. *Astérisque* **84-85**, 255–279.
- Davis, M.H.A. (1977): *Linear Estimation and Stochastic Control*. London, Chapman and Hall.
- Dietz, H.M. (1992): A non-Markovian relative of the Ornstein-Uhlenbeck process and some of its local statistical properties. *Scand. J. Statist.* **19**, 363–379.
- Dzhaparidze, K. and Spreij, P. (1993): The strong law of large numbers for martingales with deterministic quadratic variation. *Stochastics and Stochastic Reports* **42**, 53–65.
- Feigin, P.D. (1985): Stable convergence of semimartingales. *Stoch. Proc. Appl.* **19**, 125–134.
- Gushchin, A.A. and Küchler, U. (1997): Asymptotic properties of maximum likelihood estimators for a class of linear stochastic equations with time delay. Preprint, Humboldt University of Berlin.
- Hall, P. and Heyde, C.C. (1980): *Martingale Limit Theory and Its Application*. New York, Academic Press.

Helland, I.S. (1982): Central limit theorems for martingales with discrete or continuous time. *Scand. J. Statist.* **9**, 79–94.

Hutton, J.E. and Nelson, P.I. (1984): A mixing and stable central limit theorem for continuous time martingales. Technical Report No. 42, Kansas State University.

Kaufmann, H. (1987): On the strong law of large numbers for multivariate martingales. *Stoch. Proc. Appl.* **26**, 73–85.

Le Breton, A. (1977): Parameter estimation in a linear stochastic differential equation. *Transactions of the 7th Prague Conference on Information Theory, Statistical Decisions and Random Processes*, Vol. **A**, 353–366, Academia, Prague.

Le Breton, A. (1984): Propriétés asymptotiques et estimation des paramètres pour les diffusions gaussiennes homogènes hypoelliptiques dans le cas purement explosif. *C. R. Acad. Sc. Paris, Ser. I* **299**, 185–188.

Le Breton, A. and Musiela, M. (1982): Estimation des paramètres pour les diffusions gaussiennes homogènes hypoelliptiques. *C. R. Acad. Sc. Paris, Ser. I* **294**, 341–344.

Le Breton, A. and Musiela, M. (1985): Some parameter estimation problems for hypoelliptic homogeneous Gaussian diffusions. *Banach Center Publications* **16**, 337–356.

Le Breton, A. and Musiela, M. (1986): Une loi des grands nombres pour les martingales locales continues vectorielles et son application en régression linéaire stochastique. *C. R. Acad. Sc. Paris, Ser. I* **303**, 421–424.

Le Breton, A. and Musiela, M. (1987): A strong law of large numbers for vector Gaussian martingales and a statistical application in linear regression. *Statist. Probab. Lett.* **5**, 71–73.

Le Breton, A. and Musiela, M. (1989): Laws of large numbers for semi-martingales with application to stochastic regression. *Probab. Th. Rel.*

*Fields* **81**, 275–290.

Lépingle, D. (1978): Sur le comportement asymptotique des martingales locales. In Dellacherie, C., Meyer, P.A. and Weil, M. (eds.): *Séminaire de Probabilités XII, Lecture Notes in Mathematics* **649**. Berlin, Springer Verlag.

Liptser, R.Sh. (1980): A strong law of large numbers for local martingales. *Stochastics* **3**, 217–228.

Melnikov, A.V. (1986): The law of large numbers for multidimensional martingales. *Soviet Math. Dokl.* **33**, 131–135.

Melnikov, A.V. and Novikov, A.A. (1990): Statistical inferences for semi-martingale regression models. In Grigelionis et al. (eds.): *Prob. Theory and Math. Stat., Proc. of the Fifth Vilnius Conference, Vol. 2*, 150–167. Utrecht, VSP.

Rényi, A. (1958): On mixing sequences of sets. *Acta Math. Acad. Sci. Hung.* **9**, 215–228.

Rényi, A. (1963): On stable sequences of events. *Sankhya, Ser. A* **25**, 293–302.

Sørensen, M. (1991): Likelihood methods for diffusions with jumps. In N.U. Prabhu and I.V. Basawa (eds.): *Statistical Inference in Stochastic Processes*, 67–105, New York, Marcel Dekker.

Stockmarr, A. (1996): Limits of autoregressive processes with a special emphasis on relations to cointegration theory. Ph.D.. Thesis, University of Copenhagen.